Semicircularity, Gaussianity and Monotonicity of Entropy

Hanne Schultz *†

Para el Grupo

Abstract

S. Artstein, K. Ball, F. Barthe, and A. Naor have shown (cf. [ABBN]) that if $(X_j)_{j=1}^{\infty}$ are i.i.d. random variables, then the entropy of $\frac{X_1+\cdots+X_n}{\sqrt{n}}$, $H\Big(\frac{X_1+\cdots+X_n}{\sqrt{n}}\Big)$, increases as n increases. The free analogue was recently proven by D. Shlyakhtenko in [Sh]. That is, if $(x_j)_{j=1}^{\infty}$ are freely independent, identically distributed, self-adjoint elements in a noncommutative probability space, then the free entropy of $\frac{x_1+\cdots+x_n}{\sqrt{n}}$, $\chi\Big(\frac{x_1+\cdots+x_n}{\sqrt{n}}\Big)$, increases as n increases. In this paper we prove that if $H(X_1) > -\infty$ ($\chi(x_1) > -\infty$, resp.), and if the entropy (the free entropy, resp.) is *not* a strictly increasing function of n, then X_1 (x_1, resp.) must be Gaussian (semicircular, resp.).

1 Introduction.

Shannon's entropy of a (classical) random variable X with Lebesgue absolutely continuous distribution $\mathrm{d}\mu_X(x) = \rho(x)\mathrm{d}x$, is given by

$$H(X) = -\int_{\mathbb{R}} \rho(x)\log\rho(x)\mathrm{d}x, \tag{1.1}$$

whenever the integral exists. If the integral does not exist, or if the distribution of X is not Lebesgue absolutely continuous, then $H(X) = -\infty$.

The entropy can also be written in terms of score functions and of Fisher information. Take a standard Gaussian random variable G such that X and G are independent. Let

$$X^{(t)} = X + \sqrt{t}G, \qquad t \geq 0,$$

and let $j(X^{(t)}) = \big(\frac{\partial}{\partial x}\big)^*(\mathbf{1}) \in L^2(\mu_{X^{(t)}})$ denote the score function of $X^{(t)}$ (cf. [Sh, Section 3]). Then

$$H(X) = \frac{1}{2}\int_0^{\infty}\Big[\frac{1}{1+t} - \|j(X^{(t)})\|_2^2\Big]\mathrm{d}t + \frac{1}{2}\log(2\pi e). \tag{1.2}$$

The quantity $\|j(X^{(t)})\|_2^2$ is called the Fisher information of $X^{(t)}$ and is denoted by $F(X^{(t)})$. Among all random variables with a given variance, the Gaussians are the (unique) ones with the smallest Fisher information and the largest entropy.

*As a student of the PhD-school OP-ALG-TOP-GEO the author is partially supported by the Danish Research Training Council.

†Partially supported by The Danish National Research Foundation.

A. J. Stam (cf. [St]) was the first to rigorously show that if X_1 and X_2 are independent random variables of the same variance, with $H(X_1), H(X_2) > -\infty$, then for all $t \in [0,1]$,

$$H(\sqrt{t}X_1 + \sqrt{1-t}X_2) \geq tH(X_1) + (1-t)H(X_2),$$

with equality iff X_1 and X_2 are Gaussian. It follows that if $(X_j)_{j=1}^{\infty}$ is a sequence of i.i.d. random variables with finite entropy, then

$$n \mapsto H\Big(\tfrac{X_1+\cdots+X_{2^n}}{2^{\frac{n}{2}}}\Big)$$

is an increasing function of n, and if it is not *strictly* increasing, then X_1 is necessarily Gaussian.

Knowing about Stam's result, it seems natural to ask whether the map

$$n \mapsto H\Big(\tfrac{X_1+\cdots+X_n}{\sqrt{n}}\Big)$$

is monotonically increasing as well, or even simpler: Is $H\Big(\frac{X_1+X_2+X_3}{\sqrt{3}}\Big) \geq H\Big(\frac{X_1+X_2}{\sqrt{2}}\Big)$? Surprisingly enough, it took more than 40 years for someone to answer these questions. Both questions were answered in the affirmative in [ABBN] in 2004.

In this paper we extend Stam's result by showing that if $H(X_1) > -\infty$ and if for some $n \in \mathbb{N}$,

$$H\Big(\tfrac{X_1+\cdots+X_{n+1}}{\sqrt{n+1}}\Big) = H\Big(\tfrac{X_1+\cdots+X_n}{\sqrt{n}}\Big),$$

then X_1 is necessarily Gaussian (Theorem 3.1).

Free entropy, which is the proper free analogue of Shannon's entropy, was defined by Voiculescu in [V1]. If x is a self-adjoint element in a finite von Neumann algebra $\mathcal{M}$ with faithful normal tracial state τ and if $\mu_x \in \mathrm{Prob}(\mathbb{R})$ denotes the distribution of x with respect to τ, then the free entropy of x, $\chi(x) \in [-\infty, \infty[$, is given by

$$\chi(x) = \int\int \log|s-t| \mathrm{d}\mu_x(s)\mathrm{d}\mu_x(t) + \frac{3}{4} + \frac{1}{2}\log(2\pi).$$

Exactly as in the classical case, $\chi(x)$ may be written in terms of the free analogue of the score function (the conjugate variable) and the free Fisher information. That is, if s is a (0,1)-semicircular element which is freely independent of x and if we let

$$x^{(t)} = x + \sqrt{t}s, \qquad t \geq 0,$$

then

$$\chi(x) = \frac{1}{2}\int_0^{\infty} \Big[\frac{1}{1+t} - \Phi(x^{(t)})\Big]\mathrm{d}t + \frac{1}{2}\log(2\pi e), \tag{1.3}$$

where $\Phi(x^{(t)})$ is the free Fisher information of $x^{(t)}$. In [V2] Voiculescu defines for a (non-scalar) self-adjoint variable y in $(\mathcal{M}, \tau)$ a derivation $\partial_y : \mathbb{C}[y] \to \mathbb{C}[y] \otimes \mathbb{C}[y]$ by

$$\partial_y(\mathbf{1}) = 0 \qquad \text{and} \qquad \partial_y(y) = \mathbf{1} \otimes \mathbf{1}.$$

Then the conjugate variable of y, if it exists, is the unique vector $\mathcal{J}(y) \in L^2(W^*(y))$ satisfying that for all $k \in \mathbb{N}$,

$$\langle \mathcal{J}(y), y^k \rangle = \langle \mathbf{1} \otimes \mathbf{1}, \partial_y(y^k) \rangle. \tag{1.4}$$

That is, $\mathcal{J}(y) = (\partial_y)^*(\mathbf{1} \otimes \mathbf{1})$. The conjugate variable is the free analogue of the score function, and the free Fisher information of y is exactly $\|\mathcal{J}(y)\|_2^2$, so that

$$\chi(x) = \frac{1}{2}\int_0^\infty \Big[\frac{1}{1+t} - \|\mathcal{J}(x^{(t)})\|_2^2\Big] \mathrm{d}t + \frac{1}{2}\log(2\pi e). \tag{1.5}$$

Note that if $\mathcal{J}(y) = y$, then the moments of y are determined by (1.4), and it is not hard to see that y is necessarily (0,1)-semicircular.

In [Sh] D. Shlyakhtenko showed that if $(x_j)_{j=1}^\infty$ are freely independent, identically distributed self-adjoint elements in $(\mathcal{M}, \tau)$, then the map

$$n \mapsto \chi\Big(\tfrac{x_1+\cdots+x_n}{\sqrt{n}}\Big)$$

is monotonically increasing in n. In fact, the method used in [Sh] applies to the classical case as well. In this paper we will dig into the proof of the inequality

$$\chi\Big(\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big) \geq \chi\Big(\tfrac{x_1+\cdots+x_n}{\sqrt{n}}\Big) \tag{1.6}$$

and find out what it means for all of the estimates obtained in the course of the proof to be equalities. We conclude that if $\chi(x_1) > -\infty$ and if (1.6) is an equality for some n, then x_1 is necessarily semicircular. With a few modifications, our method applies to the classical case as well.

Acknowledgements. I would like to thank Dimitri Shlyakhtenko for introducing me to the problem on monotonicity of entropy and for fruitful discussions about the subject. In addition, I would like to thank the UCLA Math Department, especially the functional analysis group, for their kind hospitality during my stay at UCLA.

2 The Free Case.

Recall that the $(0,1)$-*semicircle law* is the Lebesgue absolutely continuous probability measure on $\mathbb{R}$ with density

$$\mathrm{d}\sigma_{0,1}(t) = \frac{1}{2\pi}\sqrt{4-t^2}\; 1_{[-2,2]}(t)\;\mathrm{d}t.$$

More generally, for $\mu, \gamma \in \mathbb{R}$ with $\gamma > 0$, the (μ,γ)-*semicircle law* is the Lebesgue absolutely continuous probability measure on $\mathbb{R}$ with density

$$\mathrm{d}\sigma_{\mu,\gamma}(t) = \frac{1}{2\pi\gamma}\sqrt{4\gamma-(t-\mu)^2}\; 1_{[\mu-2\sqrt{\gamma},\mu+2\sqrt{\gamma}]}(t)\;\mathrm{d}t.$$

The parameters μ and γ refer to the first moment and the variance of $\sigma_{\mu,\gamma}$, respectively.

Throughout this section, $\mathcal{M}$ denotes a finite von Neumann algebra with faithful, normal, tracial state τ. We are going to prove:

2.1 Theorem. *Let $n \in \mathbb{N}$ and let $x_1, \ldots, x_{n+1}$ be freely independent, identically distributed self-adjoint elements in $(\mathcal{M}, \tau)$. Then*

$$\chi\Big(\frac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big) \geq \chi\Big(\frac{x_1+\cdots+x_n}{\sqrt{n}}\Big). \tag{2.1}$$

Moreover, if $\chi(x_1) > -\infty$, then equality holds in (2.1) iff x_1 is semicircular.

Monotonicity of free entropy was already proven in [Sh]. Likewise, most of the results stated in this section consist of two parts: An inequality which was proven in [Sh] or in [ABBN] and a second part which was proven by us.

2.2 Proposition. *Let $n \in \mathbb{N}$ and let $x_1, \dots, x_{n+1}$ be freely independent self-adjoint elements in $(\mathcal{M}, \tau)$ with $\tau(x_j) = 0$ and $\|x_j\|_2 = \|x_1\|_2$, $1 \le j \le n+1$. Let $a_1, \dots, a_{n+1} \in \mathbb{R}$ with $\sum_j a_j^2 = 1$, and let $b_1, \dots, b_{n+1} \in \mathbb{R}$ such that $\sum_j b_j \sqrt{1 - a_j^2} = 1$. Then*

$$\Phi\Big(\sum_{j=1}^{n+1} a_j x_j\Big) \le n \sum_{j=1}^{n+1} b_j^2 \Phi\Big(\tfrac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i\Big). \tag{2.2}$$

Moreover, if $\Phi(\sum_{i \neq j} a_i x_i)$ is finite for all j, then equality in (2.2) implies that

$$\mathcal{J}\Big(\tfrac{1}{\|x_1\|_2} \sum_{j=1}^{n+1} a_j x_j\Big) = \tfrac{1}{\|x_1\|_2} \sum_{j=1}^{n+1} a_j x_j, \tag{2.3}$$

so that $\sum_{j=1}^{n+1} a_j x_j$ is $(0, \|x_1\|_2^2)$-semicircular.

2.3 Lemma. *Let $P_1, \dots, P_m$ be commuting projections on a Hilbert space $\mathcal{H}$. If $\xi_1, \dots, \xi_m \in \mathcal{H}$ satisfy that for all $1 \le i \le m$,*

$$P_1 P_2 \cdots P_m \xi_i = 0,$$

then

$$\|P_1 \xi_1 + \ldots + P_m \xi_m\|^2 \le (m-1) \sum_{i=1}^{m} \|\xi_i\|^2. \tag{2.4}$$

Moreover, if equality holds in (2.4), then $\xi_i \in \bigoplus_{j \neq i} \mathcal{H}_j$, where

$$\mathcal{H}_j := \{\xi \in \mathcal{H} \,|\, P_k \xi = \xi,\ k \neq j,\ P_j \xi = 0\} = \Big(\bigcap_{k \neq j} P_k(\mathcal{H})\Big) \bigcap P_j^{\perp}(\mathcal{H}).$$

Proof. The inequality (2.4) is the content of [ABBN, Lemma 5]. The starting point of their proof is to write each ξ_i as an orthogonal sum,

$$\xi_i = \sum_{\varepsilon \in \{0,1\}^m \setminus (1,1,\dots,1)} \xi_\varepsilon^i,$$

where for $\varepsilon \in \{0,1\}^m \setminus (1,1,\dots,1)$,

$$\xi_\varepsilon^i \in \mathcal{H}_\varepsilon := \{\xi \in \mathcal{H} \,|\, P_j \xi = \varepsilon_j \xi,\ 1 \le j \le m\}.$$

Then

$$P_1 \xi_1 + \ldots + P_m \xi_m = \sum_{\varepsilon \in \{0,1\}^m \setminus (1,1,\dots,1)} \sum_{\varepsilon_i = 1} P_i \xi_\varepsilon^i,$$

and

$$\|P_1 \xi_1 + \ldots + P_m \xi_m\|^2 = \sum_{\varepsilon \in \{0,1\}^m \setminus (1,1,\dots,1)} \Big\| \sum_{\varepsilon_i = 1} P_i \xi_\varepsilon^i \Big\|^2.$$

For fixed $\varepsilon \neq (1,1,\ldots,1)$ there can be at most $m-1$ i's for which $\varepsilon_i = 1$. Thus, by the Cauchy-Schwarz inequality,

$$\Big\| \sum_{\varepsilon_i=1} P_i \xi_\varepsilon^i \Big\|^2 \leq \Big(\sum_{\varepsilon_i=1} \|P_i\xi_\varepsilon^i\| \Big)^2 \leq (m-1) \sum_{\varepsilon_i=1} \|P_i\xi_\varepsilon^i\|^2, \tag{2.5}$$

with the second inequality being an equality iff the vector $(\|P_i\xi_\varepsilon^i\|)_{\varepsilon_i=1}$ $(= (\|\xi_\varepsilon^i\|)_{\varepsilon_i=1})$ has $m-1$ coordinates and is parallel to the vector $v = (1,1,\ldots,1) \in \mathbb{R}^{m-1}$. In particular, if the second inequality in (2.5) is an equality for some $\varepsilon \in \{0,1\}^m$ with more than one coordinate which is zero, then $(\|P_i\xi_\varepsilon^i\|)_{i=1}^m$ must consist of zeros only. It follows now that

$$\begin{aligned} \|P_1\xi_1 + \ldots + P_m\xi_m\|^2 &\leq (m-1) \sum_{\varepsilon\in\{0,1\}^m\setminus(1,1,\ldots,1)} \sum_{\varepsilon_i=1} \|P_i\xi_\varepsilon^i\|^2 &(2.6)\\ &= (m-1) \sum_{\varepsilon\in\{0,1\}^m\setminus(1,1,\ldots,1)} \sum_{i=1}^m \|P_i\xi_\varepsilon^i\|^2 &(2.7)\\ &\leq (m-1) \sum_{i=1}^m \sum_{\varepsilon\in\{0,1\}^m\setminus(1,1,\ldots,1)} \|\xi_\varepsilon^i\|^2 &(2.8)\\ &= (m-1) \sum_{i=1}^m \|\xi_i\|^2. &(2.9) \end{aligned}$$

Moreover, equality in (2.4) implies that all the inequalities (2.5), (2.6) and (2.8) are equalities. Hence,

(i) $\xi_\varepsilon^i = P_i\xi_\varepsilon^i$ for all $\varepsilon \neq (1,1,\ldots,1)$ and all $1 \leq i \leq m$ (cf. (2.7) and (2.8)), and

(ii) by the Cauchy-Schwarz argument, for all $\varepsilon \in \{0,1\}^m$ with more than one coordinate which is zero, $\|\xi_\varepsilon^i\| \overset{(i)}{=} \|P_i\xi_\varepsilon^i\| = 0$ for all i.

Thus, if equality holds in (2.4), then $\xi_i \in P_i(\mathcal{H})$ and $\xi_i \in \bigoplus_{j\neq i} \mathcal{H}_j$, as claimed. ∎

Proof of Proposition 2.2. (2.2) is the content of [Sh, Lemma 2]. Now, assume that equality holds in (2.2) and that $\Phi(\sum_{i\neq j} a_i x_i)$ is finite for all j. We are going to "backtrack" the proof of [Sh, Lemma 2] to show that (2.3) holds. We will assume that $\|x_j\|_2 = 1$ for all j.

With

$$\xi_j = b_j \mathcal{J}\Bigg(\frac{1}{\sqrt{1-a_j^2}} \sum_{i\neq j} a_i x_i \Bigg), \qquad 1 \leq j \leq n+1,$$

equality in (2.2) implies (cf. [Sh, proof of Lemma 2]) that

$$\Phi\Big(\sum_{j=1}^{n+1} a_j x_j \Big) = \Big\| \sum_{j=1}^{n+1} \xi_j \Big\|_2^2 = n \sum_{j=1}^{n+1} \|\xi_j\|_2^2. \tag{2.10}$$

Let $M = W^*(x_1,\ldots,x_{n+1})$. We now apply Lemma 2.3 to the projections $E_1,\ldots,E_{n+1} \in B(L^2(M))$ introduced in [Sh, proof of Lemma 2]. That is, E_j is the projection onto $L^2(W^*(x_1,\ldots,\hat{x_j},\ldots,x_{n+1}))$. Note that the subspace $\mathcal{H}_j$ defined in Lemma 2.3,

$$\mathcal{H}_j = \{\xi \in L^2(M) \,|\, E_k\xi = \xi,\ k\neq j\,, E_j\xi = 0\},$$

is in this case exactly $L^2(W^*(x_j))$. Thus, the second identity in (2.10) and the fact that $\xi_j \perp \mathbb{C}\mathbf{1}$, implies that

$$\xi_j \in \bigoplus_{i \neq j} (L^2(W^*(x_i)) \ominus \mathbb{C}\mathbf{1}). \tag{2.11}$$

With $E : L^2(M) \to L^2(M)$ the projection onto $L^2(W^*(\sum_j a_j x_j))$ we have (cf. [Sh, proof of Lemma 2]):

$$\mathcal{J}\Big(\sum_{j=1}^{n+1} a_j x_j\Big) = E\Big(\sum_{j=1}^{n+1} \xi_j\Big). \tag{2.12}$$

The first identity in (2.10) then implies that $E\Big(\sum_{j=1}^{n+1} \xi_j\Big) = \sum_{j=1}^{n+1} \xi_j$, and so

$$\mathcal{J}\Big(\sum_{j=1}^{n+1} a_j x_j\Big) = \sum_{j=1}^{n+1} \xi_j \in \bigoplus_{i=1}^{n+1} (L^2(W^*(x_i)) \ominus \mathbb{C}\mathbf{1}).$$

Now choose elements $\eta_j \in L^2(W^*(x_j)) \ominus \mathbb{C}\mathbf{1}$, $1 \le j \le n+1$, such that

$$\mathcal{J}\Big(\sum_{j=1}^{n+1} a_j x_j\Big) = \sum_{j=1}^{n+1} \eta_j. \tag{2.13}$$

Then

$$0 = \Big[\sum_{i=1}^{n+1} a_i x_i, \sum_{j=1}^{n+1} \eta_j\Big] = \sum_{i \neq j} \big(a_i x_i \eta_j - \eta_i a_j x_j\big).$$

A standard application of freeness shows that for $(i,j) \neq (k,l)$, the terms $a_i x_i \eta_j - \eta_i a_j x_j$ and $a_k x_k \eta_l - \eta_k a_l x_l$ are perpendicular elements of $L^2(M)$. Thus, the above identity implies that for all $i \neq j$,

$$a_i x_i \eta_j = a_j \eta_i x_j. \tag{2.14}$$

With $L^2(W^*(x_j))^0 = L^2(W^*(x_j)) \ominus \mathbb{C}\mathbf{1}$, $1 \le j \le n+1$, consider the free product of Hilbert spaces

$$\mathbb{C}\mathbf{1} \oplus \Bigg(\bigoplus_{p \ge 1} \Big(\bigoplus_{1 \le i_1, \ldots, i_p \le n+1,\ i_1 \neq i_2 \neq \cdots \neq i_p} L^2(W^*(x_{i_1}))^0 \otimes L^2(W^*(x_{i_2}))^0 \otimes \cdots \otimes L^2(W^*(x_{i_p}))^0\Big)\Bigg),$$

and notice that $x_i \in L^2(W^*(x_i))^0$ and $\eta_j \in L^2(W^*(x_j))^0$. It follows from unique decomposition within the free product that there is only one way that (2.14) can be fulfilled, namely when η_j is proportional to x_j. That is, there exist $c_1, \ldots, c_{n+1} \in \mathbb{R}$ such that $\eta_j = c_j x_j$ and hence,

$$\mathcal{J}\Big(\sum_{j=1}^{n+1} a_j x_j\Big) = \sum_{j=1}^{n+1} c_j x_j. \tag{2.15}$$

We can assume that $a_1, \ldots, a_{n+1} > 0$, and then by (2.14),

$$c_j = \frac{c_1 a_j}{a_1}, \qquad 1 \le j \le n+1.$$

In particular, all the c_j's have the same sign. Taking inner product with $\sum_{j=1}^{n+1} a_j x_j$ in (2.15), we find that

$$\sum_{j=1}^{n+1} a_j c_j = 1, \tag{2.16}$$

so that the c_j's must be positive. Also, since $\sum_j a_j^2 = 1$, we have that $\sum_j c_j^2 \geq 1$. But

$$\sum_{j=1}^{n+1} c_j^2 = \frac{c_1^2}{a_1^2},$$

and so $c_1 \geq a_1$, and in general, $c_j \geq a_j$. Then by (2.16), $c_j = a_j$, and (2.3) holds. As mentioned in the introduction, this implies that $\sum_{j=1}^{n+1} a_j x_j$ is (0,1)-semicircular (when $\|x_1\|_2 = 1$). ∎

2.4 Corollary. *Let $x_1, \ldots, x_{n+1}$ be as in Proposition 2.2 and let $a_1, \ldots, a_{n+1} \in \mathbb{R}$ with $\sum_j a_j^2 = 1$. Then*

$$\chi\Big(\sum_{j=1}^{n+1} a_j x_j\Big) \geq \sum_{j=1}^{n+1} \frac{1-a_j^2}{n} \chi\bigg(\frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i\bigg). \tag{2.17}$$

Moreover, if $\chi(\sum_{i \neq j} a_i x_i) > -\infty$ for all j, then equality in (2.17) implies that $\sum_j a_j x_j$ is semicircular.

Proof. The inequality (2.17) was proven by D. Shlyakhtenko in [Sh, Theorem 2]. Now, assume that $\chi(\sum_{i \neq j} a_i x_i) > -\infty$ for all j and that

$$\chi\Big(\sum_{j=1}^{n+1} a_j x_j\Big) = \sum_{j=1}^{n+1} \frac{1-a_j^2}{n} \chi\bigg(\frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i\bigg).$$

Take (0,1)-semicirculars $s_1, \ldots, s_{n+1}$ such that $x_1, \ldots, x_{n+1}, s_1, \ldots, s_{n+1}$ are free, and put

$$x_j^{(t)} = x_j + \sqrt{t}\, s_j.$$

Then by assumption,

$$\int_0^\infty \left[\sum_{j=1}^{n+1} \frac{1-a_j^2}{n} \Phi\bigg(\frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i^{(t)}\bigg) - \Phi\Big(\sum_{j=1}^{n+1} a_j x_j^{(t)}\Big)\right] \mathrm{d}t = 0. \tag{2.18}$$

Applying Proposition 2.2 with $b_j = \frac{1}{n}\sqrt{1-a_j^2}$, we see that the integrand in (2.18) is positive. Thus, (2.18) can only be fulfilled if for a.e. $t > 0$,

$$\sum_{j=1}^{n+1} \frac{1-a_j^2}{n} \Phi\bigg(\frac{1}{\sqrt{1-a_j^2}} \sum_{i \neq j} a_i x_i^{(t)}\bigg) = \Phi\Big(\sum_{j=1}^{n+1} a_j x_j^{(t)}\Big). \tag{2.19}$$

In fact, since both sides of (2.19) are right continuous functions of t (cf. [V2]), we have equality for all $t > 0$. Then by Proposition 2.2, $\sum_{j=1}^{n+1} a_j x_j^{(t)}$ is semicircular. By additivity of the $\mathcal{R}$-transform, this can only happen if $\sum_{j=1}^{n+1} a_j x_j$ is semicircular. ∎

Proof of Theorem 2.1. The inequality (2.1) was proven by D. Shlyakhtenko in [Sh]. Now, assume that $\chi(x_1) > -\infty$ and that

$$\chi\Big(\frac{x_1 + \cdots + x_{n+1}}{\sqrt{n+1}}\Big) = \chi\Big(\frac{x_1 + \cdots + x_n}{\sqrt{n}}\Big).$$

If we replace x_j by $\frac{x_j-\tau(x_j)}{\|x_j-\tau(x_j)\|_2}$, we will still have equality. Hence, we will assume that $\tau(x_j)=0$ and that $\|x_j\|_2=1$. Now,

$$\chi\Big(\frac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big)=\frac{1}{n+1}\sum_{j=1}^{n+1}\chi\Big(\frac{1}{\sqrt{n}}\sum_{i\neq j}x_i\Big),$$

and by application of Corollary 2.4 with $a_j=\frac{1}{\sqrt{n+1}}$, $\frac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}$ must be semicircular. Additivity of the $\mathcal{R}$-transform tells us that this can only happen if x_1 is semicircular. ∎

We would like to thank Serban Belinschi for pointing out to us the following consequence of Theorem 2.1:

2.5 Corollary. *Among the freely stable compactly supported probability measures on $\mathbb{R}$, the semicirle laws are the only ones with finite free entropy.*

Proof. By definition, a compactly supported probability measure μ on $\mathbb{R}$ is freely stable if for all $n\in\mathbb{N}$, there exist $a_n>0$, $b_n\in\mathbb{R}$, such that if $x_1,\ldots,x_n$ are freely independent self-adjoint elements which are distributed according to μ, then

$$\frac{1}{a_n}(x_1+\cdots+x_n)+b_n$$

has distribution μ. Note that the set of freely stable laws is invariant under transformations by the affine maps $(\phi_{s,r})_{s\in\mathbb{R},r>0}$, where

$$\phi_{s,r}(t)=\frac{t-s}{r},\qquad t\in\mathbb{R}.$$

Also, by [VDN, p. 27], the semicirle laws are freely stable.

Suppose now that μ is a freely stable compactly supported probability measure on $\mathbb{R}$. By the above remarks, we can assume that μ has first moment 0 and variance 1.

Let x_1,x_2 be freely independent self-adjoint elements in distributed according to μ. Since μ is freely stable, $\frac{x_1+x_2}{\sqrt{2}}$ has distribution μ as well (by the assumptions on μ, $a_2=\sqrt{2}$ and $b_2=0$). But then

$$\chi\Big(\frac{x_1+x_2}{\sqrt{2}}\Big)=\chi(x_1),$$

and by Theorem 2.1, either $\chi(x_1)=-\infty$, or x_1 is semicircular. ∎

3 The Classical Case.

In this section we are going to prove the classical analogue of Theorem 2.1:

3.1 Theorem. *Let $n\in\mathbb{N}$, and let $X_1,\ldots,X_{n+1}$ be i.i.d. random variables. Then*

$$H\left(\frac{X_1+\cdots+X_{n+1}}{\sqrt{n+1}}\right)\geq H\left(\frac{X_1+\cdots+X_n}{\sqrt{n}}\right). \tag{3.1}$$

Moreover, if $H(X_1)>-\infty$ and if (3.1) *is an equality, then X_1 is Gaussian.*

3.2 Lemma. *Let $n \in \mathbb{N}$. Then for every $m \in \mathbb{N}$, the m'th Hermite polynomial, H_m, satisfies:*

$$n^{\frac{m}{2}} H_m\Big(\tfrac{x_1+\cdots+x_n}{\sqrt{n}}\Big) = \sum_{k_1,\dots,k_n \geq 0,\, \sum_j k_j = m} \frac{m!}{k_1! k_2! \cdots k_n!} H_{k_1}(x_1) H_{k_2}(x_2) \cdots H_{k_n}(x_n). \tag{3.2}$$

Sketch of proof. (3.2) holds for $m = 0$ ($H_0 = 1$) and for $m = 1$ ($H_1(x) = 2x$). Now, for general $m \in \mathbb{N}$,

$$H_{m+1}(x) = 2xH_m(x) - 2mH_{m-1}(x).$$

(3.2) then follows by induction over m. ■

3.3 Lemma. *Let $\mu \in \mathrm{Prob}(\mathbb{R})$ be absolutely continuous w.r.t. Lebesgue measure, and let $\sigma_t \in \mathrm{Prob}(\mathbb{R})$ denote the Gaussian distribution with mean 0 and variance t. Then if $\mu((-\infty, 0]) \neq 0$ and $\mu([0, \infty)) \neq 0$, the following inclusion holds:*

$$L^2(\mathbb{R}, \mu * \sigma_t) \subseteq L^2(\mathbb{R}, \sigma_t). \tag{3.3}$$

Proof. Let $f \in L^1(\mathbb{R})$ denote the density of μ w.r.t. Lebesgue measure. Then the density of $\mu * \sigma_t$ is given by

$$\begin{aligned} \frac{\mathrm{d}(\mu * \sigma_t)}{\mathrm{d}s}(s) &= \frac{1}{\sqrt{2\pi t}} \left(\int_{-\infty}^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} \cdot e^{\frac{su}{t}} \mathrm{d}u \right) \cdot e^{-\frac{s^2}{2t}} \\ &= \phi(s) \cdot \frac{\mathrm{d}\sigma_t}{\mathrm{d}s}(s), \end{aligned}$$

where

$$\phi(s) = \int_{-\infty}^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} \cdot e^{\frac{su}{t}} \mathrm{d}u. \tag{3.4}$$

It follows that if ϕ is bounded away from 0, then (3.3) holds. For $s \geq 0$ we have that

$$\begin{aligned} \phi(s) &\geq \int_0^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} \cdot e^{\frac{su}{t}} \mathrm{d}u \\ &\geq \int_0^{\infty} f(u) \cdot e^{-\frac{u^2}{2t}} \mathrm{d}u, \end{aligned}$$

and similarly for $s \leq 0$:

$$\phi(s) \geq \int_{-\infty}^{0} f(u) \cdot e^{-\frac{u^2}{2t}} \mathrm{d}u.$$

Since $\int_{-\infty}^0 f(u)\mathrm{d}u > 0$ and $\int_0^\infty f(u)\mathrm{d}u > 0$, both of the integrals $\int_0^\infty f(u) \cdot e^{-\frac{u^2}{2t}} \mathrm{d}u$ and $\int_{-\infty}^0 f(u) \cdot e^{-\frac{u^2}{2t}} \mathrm{d}u$ are strictly positive. This completes the proof. ■

Proof of Theorem 3.1. The inequality (3.1) was proven in [ABBN]. Now, suppose $H(X_1) > -\infty$ and that (3.1) is an equality. We can assume that X_1 has first moment 0 and second moment 1. Take Gaussian random variables $G_1, \dots, G_{n+1}$ of mean 0 and variance 1 such that $X_1, \dots, X_{n+1}, G_1, \dots, G_n, G_{n+1}$ are independent. Then with

$$X_j^{(t)} = X_j + \sqrt{t}\, G_j,$$

$$H\Big(\tfrac{X_1+\cdots+X_{n+1}}{\sqrt{n+1}}\Big)=\frac{1}{2}\int_0^\infty\left[\frac{1}{1+t}-\Big\|j\Big(\tfrac{X_1^{(t)}+\cdots+X_{n+1}^{(t)}}{\sqrt{n+1}}\Big)\Big\|_2^2\right]\mathrm{d}t+\frac{1}{2}\log(2\pi e), \tag{3.5}$$

where

$$j\Big(\tfrac{X_1^{(t)}+\cdots+X_{n+1}^{(t)}}{\sqrt{n+1}}\Big)=\Big(\frac{\mathrm{d}}{\mathrm{d}x}\Big)^*(\mathbf{1})\in L^2\Bigg(\mathbb{R},\mu_{\frac{X_1^{(t)}+\cdots+X_{n+1}^{(t)}}{\sqrt{n+1}}}\Bigg) \tag{3.6}$$

is the score function. Since X_1 has mean 0 and finite entropy, μ_{X_1} and $\mu_{\frac{X_1+\cdots+X_{n+1}}{\sqrt{n+1}}}$ satisfy the conditions of Lemma 3.3.

For $t>0$, define $f^{(t)}\in L^2(\mathbb{R}^{n+1},\otimes_{j=1}^{n+1}\mu_{X_j^{(t)}})$ by

$$f^{(t)}(x_1,\ldots,x_{n+1})=j\Big(\tfrac{X_1^{(t)}+\cdots+X_{n+1}^{(t)}}{\sqrt{n+1}}\Big)\Big(\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big).$$

As in the free case (cf. (2.13)) equality in (3.1) implies that for each $t>0$ there exists a function $g^{(t)}\in L^2(\mu_{X_1^{(t)}})$ such that $\int g^{(t)}\mathrm{d}\mu_{X_1^{(t)}}=0$ and

$$f^{(t)}(x_1,\ldots,x_{n+1})=\sum_{j=1}^{n+1}g^{(t)}(x_j). \tag{3.7}$$

Because of Lemma 3.3 we can now write things in terms of the Hermite polynomials $(H_m)_{m=0}^\infty$. That is, there exist scalars $(\alpha_m)_{m=1}^\infty$ and $(\beta_m)_{m=1}^\infty$ such that

$$f^{(1)}(x_1,\ldots,x_{n+1})=\sum_{m=1}^\infty\alpha_m H_m\Big(\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big),$$

and

$$g^{(1)}(x)=\sum_{m=1}^\infty\beta_m H_m(x).$$

By Lemma 3.2, this implies that

$$\begin{aligned}&\sum_{j=1}^{n+1}\sum_{m=1}^\infty\beta_m H_m(x_j)=\\ &\quad\sum_{m=1}^\infty\frac{\alpha_m}{(n+1)^{\frac{m}{2}}}\sum_{\substack{k_1,\ldots,k_{n+1}\ge 0,\\ \sum_j k_j=m}}\frac{m!}{k_1!k_2!\cdots k_{n+1}!}H_{k_1}(x_1)H_{k_2}(x_2)\cdots H_{k_{n+1}}(x_{n+1}).\end{aligned} \tag{3.8}$$

The functions $(H_{k_1}(x_1)H_{k_2}(x_2)\cdots H_{k_{n+1}}(x_{n+1}))_{k_1,\ldots,k_{n+1}\ge 0}$ are mutually perpendicular in $L^2(R^{n+1},\otimes_{j=1}^{n+1}\sigma_1)$. Fix $m\ge 2$, and take $k_1,\ldots,k_{n+1}$ with $\sum_j k_j=m$ and $k_j\ge 1$ for at least two j's. Then take inner product with $H_{k_1}(x_1)H_{k_2}(x_2)\cdots H_{k_{n+1}}(x_{n+1})$ on both sides of (3.8) to see that α_m must be zero. That is,

$$j\Big(\tfrac{X_1^{(1)}+\cdots+X_{n+1}^{(1)}}{\sqrt{n+1}}\Big)\Big(\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big)=\alpha_1 H_1\Big(\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big)=2\alpha_1\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}.$$

Since the score function of a random variable X, $j(X)$, satisfies $\langle j(X),X\rangle_{L^2(\mu_X)}=1$, we have that $\alpha_1=\frac{1}{2}$, and so

$$j\Big(\tfrac{X_1^{(1)}+\cdots+X_{n+1}^{(1)}}{\sqrt{n+1}}\Big)\Big(\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}\Big)=\tfrac{x_1+\cdots+x_{n+1}}{\sqrt{n+1}}.$$

Then $\frac{X_1^{(1)}+\cdots+X_{n+1}^{(1)}}{\sqrt{n+1}}$ has Fisher information 1, implying that it is standard Gaussian. As in the free case, using additivity of the logarithm of the Fourier transform, this can only happen if X_1 is Gaussian. ■

References

[ABBN] S. Artstein, K. Ball, F. Barthe, A. Naor, 'Solution of Shannon's problem on monotonicity of entropy', *Journal Amer. Math. Soc.* **17** (2004), 975–982.

[Sh] D. Shlyakhtenko, 'A free analogue of Shannon's problem on monotonicity of entropy'. Preprint at http://xxx.lanl.gov/abs/math.OA/0510103.

[St] A. J. Stam, 'Some inequalities satisfied by the quantities of information of Fisher and Shannon', *Information & Control* **2** (1959), 101–112.

[VDN] D. V. Voiculescu, K. Dykema, A. Nica, 'Free Random Variables', CMR Monograph Series 1, *American Mathematical Society* (1992).

[V1] D. V. Voiculescu, 'The analogues of entropy and of Fisher's information measure in free probability theory I', *Comm. Math. Phys.* **155** (1993), 71–92.

[V2] D. V. Voiculescu, 'The analogues of entropy and of Fisher's information measure in free probability theory V', *Invent. Math.* **132** (1998), 189–227.

[V3] D. V. Voiculescu, 'Free entropy', *Bull. London Math. Soc.* **34** (2002), no. 3, 257–278.

Hanne Schultz,
Department of Mathematics and Computer Science,
University of Southern Denmark,
Campusvej 55,
5230 Odense M, Denmark
schultz@imada.sdu.dk